MAMMALIAN TOLLOID-LIKE GENE AND PROTEIN

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of Provisional Patent Application Number 60/018,684 filed May 30, 1996.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

Not applicable.

BACKGROUND OF THE INVENTION

The present invention relates to the field of bone morphogenetic proteins and more particularly to a gene in the BMP-1/Tld family of genes.

Bone formation in mammals such as mice and humans is governed by a set of bone morphogenetic proteins (BMP). Of the seven BMPs known to participate in osteogenesis, six (designated BMP-2 through BMP-7) belong to the TGF-β super family. The seventh BMP (designated BMP-1) is not TGF-β-like, but instead appears to derive from a different gene family. The BMP-1 gene family members typically contain the following domains: an astacin-like metalloprotease domain, one or more EGF-like motifs which in other proteins are thought to bind Ca^{++}, and a number of CUB domains. A CUB domain is a motif that mediates protein-protein interactions in complement components C1r/C1s which has also been identified in various proteins involved in developmental processes. BMP-1 was described, at the nucleotide sequence level, by Wozney, J.M., et al., Science 242:1528-1534 (1988).

The mammalian BMP-1 domain structure is shared by proteins found in other non-mammalian species. These proteins include *Drosophila* tolloid (Tld) (Shimell, M.J., Cell 67:469-481 (1991)), a tolloid-like *Drosophila* gene product (Tlr-1 or tolkin) (Nguyen, T., Dev. Biol. 166:569-586 (1994) and Finelli, A.L., et al., Genetics 141:271-281 (1995)), a sea urchin BMP-1

homolog (suBMP-1) (Hwang, S.-P., et al., Development 120:559-568 (1994)), two related sea urchin developmental gene products, SpAN and BP10 (Reynolds, S.D., et al., Development 114:769-786 (1992) and Lepage, T., et al., Development 114:147-164 (1992)), a *Xenopus* BMP-1 (xBMP-1) (Maeno, M. et al., Gene 134:257-261 (1993) and a mammalian tolloid (mTld) (Takahara, K. et al., J. Biol. Chem. 269:32572-32578 (1994)). A tolloid-like gene (xolloid) obtained from *Xenopus* has been briefly mentioned in passing in a article reviewing the astacin family of metalloproteases. Bond, J.S. and R. J. Benynon, Protein Science 4:1247-1261 at 1249 (1995), but data relating to the gene itself has not been published. Some of the nucleic acid sequences of the genes that encode these proteins are known. The mammalian BMP1 gene encodes both the BMP-1 protein and the mTld protein, albeit on two distinct, alternately spliced mRNA molecules. The papers mentioned in this paragraph are incorporated herein by reference.

BRIEF SUMMARY OF THE INVENTION

The present invention is summarized in that a novel mammalian tolloid-like gene product (mTll) and its cognate gene, which is distinct from mTld and from all other known BMP-1-related proteins and genes, are described. The murine and human versions of the gene are reported.

It is an object of the present invention to provide a gene and gene product involved in the deposition of extracellular matrix in vertebrates (~~or involved~~ in osteogenesis).

It is another object of the present invention to provide a target molecule for rational development of a drug for inhibiting activity of the tolloid-like genes to treat fibrosis, scarring, keloids, surgical adhesions, and the like.

It is yet another object of the present invention to provide a recombinant DNA construct, and a protein encoded by the construct, for use in accelerated wound and fracture healing.

It is still another object of the present invention to provide a marker gene that maps to the central portion of mouse

chromosome 8.

It is yet another object of the present invention to provide a marker gene that maps to the 4q32-4q33 region of human chromosome 4.

It is still another object of the present invention to provide a nucleotide sequence that functions as a probe for a non-BMP-1 bone morphogenetic protein gene in mammalian cells.

It is a feature of the present invention that the murine gene described contains a novel simple sequence repeat in the 3'-untranslated region of the gene.

Other objects, features, and advantages of the present invention will become apparent upon consideration of the following detailed description considered in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

Figure 1 presents a map of the murine mTll cDNA including 5'- and 3'-untranslated portions thereof. Also aligned beneath the mTll cDNA for comparison are schematic representations of cDNA clones of related genes, drawn to the same scale as mTll. Portions of the cDNA corresponding to domains of the gene product are highlighted. Stippled, darkly shaded, striped, lightly shaded, and black boxes represent signal peptide, proregion, metalloprotease, CUB, and EGF domains, respectively. White boxes represent domains unique to the various proteins. Wavy lines represent 5'- and 3'-untranslated regions. Abbreviations: mTld, mammalian tolloid; mBMP-1, mammalian BMP-1; xBMP-1, xenopus BMP-1; suBMP-1, sea urchin BMP-1; Tld, *Drosophila* tolloid; Tlr-1, *Drosophila* tolloid-related gene; SpAN and BP10, related sea urchin developmental genes. Restriction enzymes include: Bg, BglII; C, ClaI; E, EcoRI; H, HincII; N, NcoI; S, SmaI; St, StuI.

Figure 2 aligns the amino acid sequence of the disclosed mTll gene to that of the mTld gene. The domain structure common to both proteins is shown schematically. Domains are represented as in Figure 1. Alignment was performed using the GAP program (Genetics Computer Group, Madison, Wisconsin), with

a GAP weight of 3.0 and GAP length weight of 0.1, with some additional manual alignment of putative signal peptide sequences. Cysteines are boxed, potential Asn-linked glycosylation sites are underlined and the metalloendopeptidase active site motif HEXXH is enclosed by a dashed box.

Figure 3 shows a schematic map of the central portion of mouse chromosome 8. The Map Manager program (Manly K., "A Macintosh Program for Storage and Analysis of Experimental Genetic Mapping Data," Mammalian Genome 4:301-313 (1993) compared segregation data for Tll and for other loci from the TJL BSS backcross panel, performed the linkage analysis and generated the map. The TJL BSS backcross panel data are available on the Jackson Laboratories Public Data Base (http://www.jax.org/Resources/documents/cmdata).

DETAILED DESCRIPTION OF THE INVENTION

A substantially pure preparation of the mammalian tolloid-like (mTll) cDNA was isolated from mice by probing a cDNA library prepared from embryo fibroblasts of mouse strain BMP-1^{tm1b1h} with an approximately 330 base-pair AatII-DrdI restriction fragment of the mouse BMP-1 gene and screening at low stringency, as is described in more detail in the examples below. BMP-1^{tm1b1h} is a BMP-1 knockout (KO) mouse that is homozygous for a null allele of the BMP-1 gene. The probe, shown in SEQ ID NO:1, corresponds to a segment of the 360 base pair portion of the BMP-1 gene that is absent from the BMP-1^{tm1b1h} knockout strain. Since BMP-1 is absent from the cDNA library, the screen uncovered only sequences related to, but distinct from, BMP-1 at the DNA sequence level.

A single region of the genome was uncovered in the screening (see Fig. 1). Overlapping cDNA clones KO 3 and KO 7-2, obtained in the initial screen as substantially pure preparations covered much, but not all, of the coding sequence. Screening was performed under low stringency using standard protocols (Ausubel, F.M., et al., Current Protocols in Molecular Biology, Wiley, New York (1987)CA). The remainder of the coding sequence was obtained by re-screening the embryo

fibroblast cDNA library at high stringency to reveal an additional clone (KO 8-2) that extended into the 5'-untranslated portion of the gene. Clone KO 8-2 was also obtained as a substantially pure preparation.

One of ordinary skill in the art can join the separate cloned sequences together, as the inventors have done, to produce the complete full-length cDNA shown in SEQ ID NO:2. Presented herein as SEQ ID NO:2 is an open reading frame of 3039 base pairs flanked by 5'- and 3'- untranslated sequences. The open reading frame encodes a mammalian tolloid-like protein termed mTll. The sequence presented herein represents the combined nucleic acid sequences of three cDNA clones (KO 3, KO 7-2 and KO 8-2). In the 3' untranslated region of the mTll gene is a previously unreported simple sequence repeat (SSR). The SSR has the following sequence: $(GT)_{20}GC(GT)_7$ $GC(GT)_7GCAT(GT)_3GCAT(GT)_3$ (shown at nucleotides 4148 to 4239 in SEQ ID NO:2). The sequence data presented in SEQ ID NO:2 will be available at Genbank Accession Number U34042..

A preparation of DNA molecules containing an mTll gene sequence from any source is considered substantially pure if more than 90% of any cellular material of any host cell in which the DNA has resided has been removed from the DNA preparation. Cellular material can be removed from a nucleic acid preparation, for example, by using a commercial purification kit such as is available from Qiagen (Chatsworth, CA). It is preferred that greater than 10% of the nucleic acid molecules in a nucleic acid preparation comprise the complete or partial mTll gene or a portion thereof. More preferably, greater than 50%, and yet more preferably, greater than 90%, of the nucleic acid molecules comprise the complete or partial sequence.

It is noted that additional genes having some relation to mTll and other members of the BMP-1 family might be isolated from double null embryo cDNA libraries lacking both BMP-1/mTld and mTll using a comparable screening strategy. Such double null mutant animals could be produced by mating animals heterozygous at each of the two loci.

The murine mTll gene maps to central chromosome 8 close to D8Bir22, which was placed at position 31 in the 1994 chromosome 8 committee report (Ceci, J.D., "Mouse Chromosome 8," Mammalian Genome 5:S124-S138 (1994)). This same general chromosomal region is the map site of four genetic defects that lead to apparent developmental abnormalities: Hook (Hk), Adrenocortical dysplasia (Acd), Quinky (Q), and Proportional dwarf (pdw). BMP-1, in contrast, maps to mouse chromosome 14 (Ceci, J.D., et al., "An interspecific backcross linkage map of the proximal half of mouse chromosome 14," Genomics 6:673-678 (1990).

A proposed murine mTll protein domain structure, predicted from sequence similarities to the m-tolloid protein product, is shown in Figure 2. In view of the similarity to other tolloid-like proteins, it is expected that the product encoded by the disclosed mTll gene will be a protease having a key role in development and in homeostatic processes such as wound healing. It is likely that the protein is involved in maturation of extracellular matrix precursors into macromolecular structures. The protein may also have a role in activation of growth factors *in vivo* and *in vitro*, and may accelerate developmental and homeostatic processes when an effective amount of the protein is administered to a tissue. On the other hand, if the mTll protein function is inhibited, such processes may themselves be inhibited, which property can be exploited advantageously upon delivery of an effective amount of an inhibitor to prevent fibrosis and excess scarring or other abnormalities of wound healing. An effective amount of the protein to be delivered to a target site for activating developmental and homeostatic properties can readily be determined by testing a range of amounts of the protein on a selected veterinary species or on a model species having acknowledged biochemical or physiological similarity to humans. In the case of skin wound healing, for example, porcine skin is a suitable model for human skin. Likewise, an effective amount of an inhibitor of the Tll protein can also be determined. An effective amount is an amount effective upon administration to

a wound that reduces the occurrence of fibrosis, scarring or keloids compared to an untreated wound, where the assessment of fibrosis, scarring or keloids is made according to accepted clinical or veterinary standards. Such a test is preferably performed in a model system generally accepted as having relevance to human skin.

The ability to work with proteins of the BMP system has been hampered by the fact that the proteins are typically present in very small amounts in animal tissues. It is herein demonstrated (see, *infra*) that mTll, a previously unknown gene, can be cloned into a suitable expression vector containing a transcriptional promoter effective in a suitable host cell, introduced into and expressed in the suitable host cells, and purified in a native configuration, all using conventional methods. The protein thus expressed can remain inside the host cell or can be secreted to the extracellular growth medium, if a suitable signal sequence is provided on the construct. The protein can be purified from the cell or from the growth medium by conventional methods.

A suitable promoter of transcription is the baculovirus very late promoter found on vector pFASTBac1, which vector is commercially available from Gibco-BRL. Another suitable promoter is baculovirus immediate early promoter such as is found on the pAcPIE1 vector (Novagen, Madison, WI). Any other advantageous expression elements such as enhancers, terminators, and the like, as are known to the art, can be included on the suitable expression vector. A suitable host would be insect tissue culture cells, such as cell line Sf21, Sf9, or High Five (Invitrogen, San Diego, CA).

Suitable portions of the gene comprising less than the full coding sequence can also be advantageously cloned into the suitable expression vector to form a recombinant genetic construct. It is understood that a construct prepared in accordance with the invention, need not necessarily contain the entire mTll locus or coding region, but could contain one or more portions thereof encoding a desired function, or containing a portion of the gene having other useful

properties, for example, complementarity to a desired genomic sequence. It is understood by those of ordinary skill that certain variation in the size or sequence of the mTll protein (and in the corresponding genetic material encoding the mTll protein) will not interfere with the functions thereof. Such modified forms can be engineered using known methods that may be advantageously employed when constructing genetic constructs containing the complete or partial mTll gene, and in proteins encoded thereby.

Such changes, modifications, additions and deletions are contemplated to fall within the scope of the present invention, as long as the protein retains a desired function known to be associated with other members of this protein family. The protein is competent if it retains an ability to cleave laminin-5 in a standard assay for such cleavage. It is also desired that the protein retain a C-proteinase activity against procollagen as was described for BMP-1 by Kessler, E., et al., Science 271:360-362 (1996), incorporated herein by reference. One of ordinary skill is familiar with the necessary controls that should accompany any such assay. It may, alternatively, be desired that the protein lose a certain function as a result of such a change, and such a situation is also envisioned to be within the scope of the present invention.

A substantially pure preparation of the protein thus produced is defined as a preparation wherein the laminin-5-cleaving activity of the mTll protein is not affected by the presence of other proteins or molecules in the preparation. Depending upon the use to which the protein will be put, it may be that the mTll protein accounts for at least 10%, preferably at least 50%, more preferably at least 75%, and most preferably at least 95% of the protein in the substantially pure protein preparation. The protein preparation can be enhanced for the protein of interest by labeling the protein with an affinity tag and passing the preparation over a column having an affinity for the tag. It is also possible to employ a processing tag such that a properly processed form of the protein (lacking the cleaved proregion) can be eluted from a

column loaded with a crude preparation.

The mTll translation product (SEQ ID NO:3) predicted from the DNA sequence has a predicted molecular weight of 114,532 (pI 6.15). If the translation product is cleaved between the proregion and the protease domain at the boundary shown in Figure 2, the predicted molecular weight for the mature protease would be 98,007 (pI 6.18).

When the murine mTll protein sequence is compared to other tolloid-like genes, no obvious homology exists between the proregion of either of the two mammalian proteins (mTld and mTll) and the proregion of either of the two *Drosophila* proteins (Tld or Tlr-1). The protease domain of mTll was 66% similar (47% identical) to Tld and was 69% similar (52% identical) to Tlr-1. mTll is slightly more similar in sequence to both *Drosophila* proteins than is mTld, and there is no obvious correlation between a particular member of the mammalian protein pair and a particular member of the *Drosophila* protein pair. An aligned pair of amino acids are "similar" if they have a threshold of similarity above 0.5 by the scoring system of Schwartz and Dayhoff, Atlas of Protein Sequence and Structure, Dayhoff, M.O., ed., National Biomedical Research Foundation, Washington, D.C., p. 353-358 (1979).

The mTll mRNA transcript appears not to be alternatively spliced since only a single transcript was detected using a fragment of clone KO 3 internal to the coding region as a probe (SEQ ID NO:2, nucleotides 1113 to 2745) and because only a single mTll cDNA was isolated during the cDNA library screenings.

Relatively strong mTll mRNA expression was observed in adult brain and kidney, with somewhat lower expression in RNA from lung and skeletal muscle, and very low expression in RNA from heart and testes. No signal was apparent for spleen or liver. After the Northern Blot was exposed for 60 hours, a very faint signal could be detected for liver, although no signal from spleen was detected.

The mRNA expression pattern of mTll differs from that previously reported for BMP-1 transcripts and mTld transcripts.

Low expression levels are seen even in seven day post-coitum total embryo RNA. The mRNA level increases slightly at eleven days of development, peaks at relatively high levels at fifteen days, and then decreases in seventeen day embryos. In contrast, BMP-1 and mTld transcripts were observed at higher levels in seven day embryos than in eleven day embryos. The same blot was used to monitor the mTll, BMP-1, and mTld transcript levels.

The mTll mRNA transcripts were detected throughout embryonic development in the period of 9.5 to 15.5 days post-coitum. As was previously observed with mTld RNA, mTll signals were observed throughout the mesenchyme, with higher levels overlying areas of future bone and the ventral portion of the neural tube. A strong signal, seen in the same portion of the ventral hindbrain in which signal was previously observed for mTld, is consistent with expression of mTll in the floor plate. A regular pattern of strong expression was observed overlying the connective tissue between the developing vertebra. The high mTll signal observed in the mesenchyma of the developing lung contrasts with the absence, or very low level, of expression in liver which mirrors the relative amounts of mTll mRNA found in adult mouse lung and liver by Northern Blot analysis. In a parasagittal section of a 13.5 day post-coitum embryo, expression was observed in mesenchymal elements of the developing tongue, nasal process, and jaw and in the submucosal layer in loops of the developing intestine. mTll expression was observed overlying a developing atrioventricular valve of the heart.

A major difference between the distribution of mTll and mTld mRNA in developing mouse tissues is seen overlying the neuroepithelium in the vestibular area of the floor of the fourth ventricle of the developing brain where strong mTll expression was consistently observed, in various sections, and where neither mTld nor BMP expression has been observed. mTll RNA expression was observed, in a number of sections, to overlie the neuroepithelial lining of the ventricles and aqueduct of neonatal brain. mTll expression was also observed

overlying specific nuclei within the thalamus and the neuroepithelial lining of the lateral ventricles. In adult brain, strong mTll expression was observed in the granular layer of the cerebellum. Weaker mTll expression was also observed overlying other structures of the neonatal and adult mouse brain. Northern blot analysis of RNA from various portions of human brain has also detected relatively strong signal for mTll in the human cerebellum.

Another difference noted between the distribution of mTll mRNA and that previously described for mTld was in a developing spinal cord where mTll expression was more extensive than was previously noted for mTld, extending beyond the floor plate toward more dorsal portions of the spinal cord. In other developing tissues, the distribution of mTll and mTld transcripts appear to overlap.

It is specifically envisioned that equivalents of the mTll gene can be isolated from other species, by probing a cDNA library from cells of an appropriate species with a probe selected to include an mTll-specific portion of the described mouse gene. An mTll-specific portion of the mouse mTll gene can be obtained by comparing the nucleic acid sequence of the mouse mTll coding region to that of BMP-1/mTld and selecting a portion of the mTll gene that has no equivalent in BMP-1. To be an effective probe, the selected sequence should not contain repeat sequences that would cross-hybridize to numerous genomic sites. The probe should be at least about 200 bases long. It is recognized that the genes of the BMP-1 family are most variable in the regions that encode the proregion and the C-terminal 17 amino acids of the proteins, and it is anticipated that suitable probes can be isolated from those regions of the mTll gene. In SEQ ID NO:2, this region corresponds to the sequence shown between about bases 3599 and 3650 for the C-terminal portion and about 701 - 1051 for the proregion. Such a fragment can be converted into a probe by nick translation, end labeling, or other suitable technique known to the art. It is also understood that a desired fragment (or indeed an entire gene) can be synthesized *in vitro* using well-known

techniques available to the molecular biologist.

This has been accomplished using human source DNA. To obtain human mTll sequences, a 677 bp NdeI-Eco72I fragment of mouse mTll cDNA clone KO 3, corresponding to a portion of CUB4 and all of CUB5 and the carboxy terminus, was used to screen a human placenta genomic DNA library. Genomic clone 151-2 was isolated which contained the final three exons of the human TLL gene. A 339 bp TaqI fragment of mouse mTll cDNA KO 7-2, corresponding to a part of the proregion was then used to screen the same human genomic DNA library resulting in isolation of genomic clones 5-2 and 8, each of which contained the first, 5'-most, exon of the Tll gene. Oligonucleotide primers were synthesized corresponding to sequences in the 5'- and 3'- untranslated regions and were used with cDNA synthesized from human fetal cartilage RNA for long distance PCR amplification of the remainder of TLL coding sequences. The forward primer was 5'-TCTTGCAGTCAGTTGCTTTGCTGG-3' (SEQ ID NO:10). The reverse primer was 5'-TAGTGCGGCCGCACATTCCTTTGTGTTC-3' (SEQ ID NO:11).

The nucleic acid sequence of human mTll is shown in SEQ ID NO:4. The protein encoded by the gene is shown in SEQ ID NO:5. The gene (or portions thereof) can be used in the same ways as the murine gene, but with the additional benefit for genetic therapies, diagnoses, and the like, since there is no need to adapt the gene for use in humans, as could be the case for the mouse mTll gene.

Because defects in mTll may lead to genetic abnormalities in people, the chromosomal position of the human TLL gene was established. A 527 bp cDNA PCR product, corresponding to the last 3 exons of the human TLL gene, was hybridized to Southern blots of EcoRI-digested genomic DNA from panels of human-mouse cell hybrids. Strong hybridization to ~5.1 and 9.5 kb human bands was observed and examination of DNA from 30 hybrid lines, derived from 17 unrelated human cell lines and 4 mouse cells lines (Takahara, K., et al., J. Biol. Chem. 269: 26280-26285 (1994)), showed that the segregation of TLL correlated with the distribution of human chromosome 4. Of the cell hybrids

examined, one that retained a translocation of human chromosome 4 further localized TLL to the chromosome 4 long arm. Cell hybrid 55R16 has no intact chromosome 4 but retains the 11/4 translocation 11qter-11p13::4q25-4qter. These results localized TLL to the 4q25-4qter region. The TLL gene was independently mapped by fluorescence in situ hybridization (FISH) on human metaphase chromosome spreads by the method of Trask, B., Methods Cell Biol. 35: 1-35 (1992). Human genomic DNA clone 8, which contains the TLL first exon and has an insert size of approximately 16 kb, was labeled with digoxygenin-11-dUTP (Boehringer Mannheim) by random priming (Feinberg, A.P., and B. Vogelstein, B. Anal. Biochem. 132: 6-13 (1983)) and employed as a probe for FISH analysis. Images were obtained and analyzed as described (Takahara, K., et al., *supra*). Double fluorescent signals were found only at 4q32-4q33 in 16/18 of the metaphase spreads examined (88.8%), with double fluorescent signals found on both chromosomes of 10/18 metaphase spreads and on no other chromosome, localizing TLL to this region.

It should also be possible to use PCR to amplify a portion of a genome that corresponds to the mTll region, by selecting specific primers expected to flank the mTll gene (or any portion of the gene). Two mTll-specific portions of the gene can serve as suitable primers. It may not be effective to select primers outside the coding portion of the gene because reduced selective pressure on non-coding portions results in greater divergence between mice and other species in those regions. It is specifically noted that the genes of the BMP family from humans and model species such as the mouse are particularly sought after for their relation to human deformities (see, e.g., "The Chicken With a Duck's Feet: It's All in the Biochemical Signal," The New York Times, National Edition, p. B6 (May 21, 1996)).

It is also specifically envisioned that large quantities of the protein encoded by the mTll gene can be expressed in (or secreted from) host cells, purified to a substantially pure preparation and used in subsequent functional assays. In one

such functional assay, functional attributes of the expressed protein will be described. The protein functions are expected to include a metalloprotease activity, C-proteinase activity and laminin-5 processing activity, and an activating activity for TGF-β-like proteins, such predictions being reasonable in view of the gross structural similarity to known proteins at the domain level.

In another assay, the protein can be used to screen putative agents having inhibitory activity against the protein. Given that mTll is able to rescue BMP-1 knockout mice, it will be important for any therapeutic system that modifies or eliminates BMP-1 protein function to similarly alter the mTll protein function. Thus, any panel of such agents must be screened against mTll protein. In such an assay, all components of an assay that support mTll function can be added together, under suitable conditions of salt and pH, and combined with a panel of putative inhibitors of protein function. Using established assays of protein function (described in documents incorporated elsewhere herein by reference), it will be possible to determine whether any tested agent can inhibit protein activity, thereby making it a likely candidate for use in a therapeutic amount to inhibit fibrosis, reduce scarring, and reduce keloids. Such screening efforts are underway using related proteins from the BMP-1 family of genes. See Kessler, *supra*.

It is now also possible to embark upon a rational drug design strategy using the disclosed protein or fragments thereof. In doing so, the protein or fragments will be subjected to x-ray crystallographic analysis to determine their active sites and sites that are available for interaction with a putative therapeutic agent.

The protein encoded by BMP-1 was recently shown to cleave procollagen near the C-terminus. This C-proteinase activity, which is essential to the production of collagen, had long been thought to reside in a protein that had remained elusive. There is great commercial interest in harnessing the C-proteinase activity as a therapeutic agent in collagen-related

diseases. Since mTll appears to be the only other mammalian gene closely related to BMP-1 (on the basis of the cDNA library screening results), it is also specifically contemplated that the protein encoded by mTll will be an alternative C-proteinase and, further, that the mTll gene can be utilized in the effort to produce an alternative C-proteinase, both by incorporating the gene into a recombinant vector for *ex vivo* production of therapeutic protein, and for direct administration in a genetic therapy. The human gene has particular utility for these applications.

The invention will be better understood upon consideration of the following non-limiting Examples.

EXAMPLES

BMP-1/mTld-null Mouse Embryo cDNA Library

Mouse embryo fibroblasts (MEFs) were prepared as described (Hogan et al., "Manipulating the Mouse Embryo: A Laboratory Manual, 2nd Ed.," pp. 260-261, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1994)). Ten 150 mm plates of MEFs prepared from embryos made homozygous for null alleles of the BMP-1/mTld gene (express no BMP-1/mTld proteins) were grown to confluence (in DMEM, 10% fetal calf serum), and 3 days later were treated with 50 μg/ml ascorbate for 18 hr, harvested, and 42 μg poly(A$^+$) mRNA was isolated using a FastTrack kit (Invitrogen). A 5 μg aliquot of poly(A$^+$) was then used for synthesis of double-stranded cDNA with EcoRI ends using the SuperScript Choice System (Gibco-BRL). This cDNA was then ligated to EcoRI-cut λgt10 arms and packaged using Gigapak II Gold packaging extract (Stratagene). The 5 μg of poly (A$^+$) provided an unamplified library of ~2.2 x 10^6 PFUs. The randomly picked clones had an average insert size of ~2.9 kb.

DNA Sequence Analysis

Restriction fragments were subcloned into pBluescript II KS$^+$ and sequences were obtained from double-stranded templates by dideoxy chain termination, as described in Lee S.-T., et al., "Construction of a full-length cDNA encoding human pro-alpha 2(I) collagen and its expression in pro-alpha 2(I)-

deficient W8 rat cells," J. Biol. Chem. 263:13414-13418 (1988). Ends of subclones were sequenced using T3 and T7 primers with internal portions of subclones made accessible to sequencing by introducing deletions or using primers complementary to insert sequences. The mTll sequences reported herein were confirmed by sequencing both strands.

Polymerase Chain Reaction (PCR)

The PCR was performed with 0.2 μM of each primer in a 480 thermal cycler (Perkin-Elmer Corp.) with denaturation at 94°C for 3 min, followed by 35 cycles of 94°C/1 min, 57°C/1 min, 72°C/1.5 min, and final incubation at 72°C/8 min. Final volumes were 100 μl of 10 mM Tris-HCL, pH 8.3, 50 mM KCL, 1.5 mM $MgCl_2$, 0.01% (w/v) gelatin, 0.2 mM each dNTP, and 2.5 units of *Taq* polymerase (Perkin-Elmer Corp.).

Tissue Sections for in situ Hybridization

Tissue sections mounted on slides for *in situ* hybridization were kindly provided by G.E. Lyons (University of Wisconsin-Madison). Mouse tissues were fixed and embedded, as in Lyons et al., "The expression of myosin genes in developing skeletal muscle," J. Cell Biol. 111:1465-1476 (1990). Briefly, tissues were fixed in 4% paraformaldehyde in phosphate-buffered saline, dehydrated, and infiltrated with paraffin. Serial sections, 5-7 μm thick, were mounted on gelatinized slides. One to three sections were mounted/slide, deparaffinized in xylene, and rehydrated. Sections were digested with proteinase K, post-fixed, treated with tri-ethanolamine/acetic anhydride, washed, and dehydrated.

Probes for *in Situ* Hybridization

mTll-specific probes corresponding to portions of the 1104 bp mTll 3'- untranslated region were used for *in situ* hybridization. Since the 3'- untranslated region has no similarity to BMP-1 or mTld sequences, the probes did not cross-hybridize with BMP-1 or mTld RNA.

To ensure that the probes did not hybridize to other RNA transcripts bearing repeat sequences similar to the long SSR identified in the central portion of the mTll 3'-untranslated region (nucleotides 4148 to 4239), two separate riboprobes,

corresponding to 3'-UT sequences upstream or downstream of the SSR were prepared according to the manufacturer's conditions (Stratagene), labeled with ^{35}S-UTP (>1000 Ci/mmol/ Amersham Corp.) and combined to strengthen the *in situ* hybridization signal. Probes were hydrolyzed with alkali to a mean size of 70 bases.

For 3'-untranslated sequences downstream of the SSR, a 399 bp PCR product (SEQ ID NO:2 nucleotides 4283 to 4681) was prepared using forward primer 5'-CCAGCTTAACCTGTTCACAC-3' (SEQ ID NO:6) and reverse primer 5'-AACTCTACTTCCACTTCATC-3' (SEQ ID NO:7). The PCR product was ligated into the cloning site of the pCRII T-A vector (Invitrogen). Uniformly labeled antisense riboprobe was generated by linearizing the template at the HindIII site in the pCRII polylinker and transcribing with RNA polymerase T7. Sense control riboprobe was generated by linearizing at the XhoI site in the pCRII polylinker and transcribing with RNA polymerase SP6.

For 3'-untranslated sequences upstream of the SSR, a 420 bp PCR product (SEQ ID NO:2, nucleotides 3666 to 4085) was prepared, employing forward primer 5'-TCAGAACAGAAAGGAATGTG-3' (SEQ ID NO:8) and reverse primer 5'-GACCACTATTCCACATCACC-3' (SEQ ID NO:9), and was ligated into the cloning site of pCRII T-A. Antisense riboprobe was prepared by linearizing at the XhoI site in the pCRII polylinker and transcribing with RNA polymerase SP6, while sense control riboprobe was prepared by linearizing at the HindIII site in the pCRII polylinker and transcribing with RNA polymerase T7.

In situ hybridization and Washing Procedures

Sections were hybridized overnight at 52°C in 50% deionized formamide, 0.3 M NaCl, 20 mM Tris-HCl, pH 7.4, 5 mM EDTA, 10 mM $NaPO_4$, 10% dextran sulfate, 1 x Denhardt's solution, 50 μg/ml total yeast RNA, 25 μmol/ml thio-ATP (Boehringer-Mannheim), and 50-75,000 cpm/μl ^{35}S-labeled cRNA probe. Tissue was stringently washed at 65°C in 50% formamide, 2 x SSC, 10 mM dithiothreitol; rinsed in phosphate-buffered saline; and treated with 20 μg/ml RNase A at 37°C for 30 min. Following washes in 2 x SSC and 0.1 x SSC for 15 min at 37°C,

slides were dehydrated, dipped in Kodak NTB-2 nuclear track emulsion, and exposed for 1 week in light-tight boxes with desiccant at 4°C. Photographic development was in Kodak D-19. Slides were analyzed using light- and dark-field optics of a Zeiss Axiophot microscope.

Northern and Southern Blot Analyses

A 1,633 bp EcoRI fragment (SEQ ID NO:2, nucleotides 1113 to 2745) corresponding to the 5'-end of cDNA clone KO 3 (Fig. 1) was purified and used as a probe for Northern blot analyses. This fragment contains sequences corresponding to most of the protease domain; all of the domains CUB1, CUB2 and EGF1; and most of domain CUB3. The 399 bp PCR product described above for use in *in situ* hybridization experiments was gel purified and used as a probe in Southern blot analyses. Both probes were radiolabeled to a specific activity of 4-6 x 10^9 cpm/μg by random priming (Feinberg and Vogelstein, "A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity: Addendum," Anal. Biochem. 137:266-267 (1984)) and were hybridized to blots in QuickHyb (Stratagene) at 68°C for 1 h. Northern blots (obtained from Clontech) were washed twice in 2 x SSC, 0.1% SDS at 68°C for 10 min and then twice in 0.1 x SSC, 0.1% SDS at 68°C for 15 min. Southern blots were washed twice in 2 x SSC, 0.1% SDS at 68°C for 10 min and then twice in 0.1 x SSC, 0.1% SDS at 68°C for 20 min.

Subcloning and expression of mTll gene

The mature active forms of BMP-1, mTld and mTll are all similar in their amino acid sequences. An exception to this is the C-terminus of each protein, where no homology is observed. This uniqueness of C-terminal sequences has been put to use in producing a set of polyclonal antibodies capable of discriminating between the three protein forms. In the case of mouse mTll, the synthetic peptide Ac-CYIRYKSIRYPETMHAKN-OH, which corresponds to the final 17 amino acids of mTll, was linked to the protein carrier Keyhole Limpet Hemocyanin, suspended in saline and emulsified by mixing with an equal volume of Freund's adjuvant and injected into three to four subcutaneous dorsal sites in each of two rabbits. Bleeds for

sera were at 12 and 16 weeks after immunization and boosts. Unlike BMP-1 and mTld, for which C-terminal amino acid sequences are perfectly conserved between mouse and human, mouse and human mTll C-terminal amino acid sequences are diverged. It is perhaps because of this divergence across species that the peptide for the mouse mTll C-terminus peptide has produced 3-fold higher titers of antibodies in rabbits than have the C-terminus peptides of BMP-1 and mTld. In order to produce antibodies specific for the C-terminus of human MTll, the peptide Ac-CHIRYKSIRYPDTTHTKK-OH will be used. These antibodies have commercial utility in an assay for visualizing the production and localization of mTll protein in cells, tissues, and mammalian organisms, including, but not limited to model systems (e.g., rodents, primates, and the like) as well as humans. In view of the rapid pace at which the understanding of the bone morphogenetic proteins is advancing, the ability to distinguish individual components one from another is important, not merely from a research perspective, but in monitoring the level and distribution of BMP system components in patients having disorders of the BMP system. Such disorders could include, for example, in mice and humans, fibrotic conditions. In addition, hereditary developmental abnormalities may be due to defects in the TLL gene. Determining the role of mTll in such genetic abnormalities will be enabled by the antibody and nucleic acid probes described herein. The mTll protein is quite clearly important in the BMP system, in that it apparently substitutes well for BMP-1 in mice having null BMP-1 alleles on both chromosomes. Such mice survive the full course of gestation but develop a persistent herniation of the gut in the umbilical region. These mice die soon after birth, presumably due to the loss of the BMP-1/mTld gene. However, they show no gross derangements of pattern formation, of collagen fibril formation, or of development in general. Clearly development of this order or even collagen fibrillogenesis would not be possible without some BMP-1/mTld-like activity. We have found such an activity in mouse embryo fibroblasts from these BMP-1-null mice in the form of C-

proteinase activity. Such activity appears to be supplied by mTll and there appear to be no other closely related genes.